Abstract Title Page
Not included in page count.

Title: The Causal Effect of the School Day Schedule on Adolescents' Academic Achievement

Authors and Affiliations:
Teny M. Shapiro, Assistant Professor of Economics – Santa Clara University
Kevin M. Williams, PhD Candidate in Economics – UC Davis

Abstract Body

Limit 4 pages single-spaced.

Background / Context:

Description of prior research and its intellectual context.

> Our study looks at the causal impact of the school day schedule on student achievement. A few, but not many, studies have explored the role of school start times and schedule on academic achievement. Cortes et al. (2012) utilize the essentially random variation between students in the ordering of classes over the day at Chicago Public High Schools and find that having a class first period significantly reduces grades in that course, and that having math in first period reduces test scores in all subjects and reduces grades in future math classes as well. Dills & Hernandez-Julian (2008) find that students perform better in classes that meet later in the day. Contrary to that, Pope (2014) uses data from Los Angeles County schools to determine how secondary-school students perform in morning versus afternoon courses. He finds that learning monotonically decreases throughout the school day. Edwards (2012) studied the effect of start times on middle school students and found that a one hour delay in start times leads to a three percentile point gain in both math and reading test scores for the average student. Carrell et al. (2011) study the role of school start times at USAFA by utilizing two policy changes in the daily schedule during a three year period. They find that starting the school day 50 minutes later increases overall academic achievement by about one-tenth of a standard deviation and that performance throughout the day is affected by early start times.

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

> How a student's classes are scheduled throughout the day is often determined by necessity, but can have a meaningful impact on academic performance. Acknowledging students' internal clocks and making small changes to scheduling patters could be a relatively low-cost method for administrators to improve performance. This paper builds on literature that has shown the negative effects of early morning classes to consider the influence the school-day schedule has throughout the day. Our data is five cohorts of college freshman at the United States Air Force Academy who face randomized scheduling and largely take a common set of classes. We find the largest impact of the schedule is in the early morning, but also find evidence of academic fatigue and asymmetric effects among STEM, non-STEM, and physical education courses that vary over the day. Contrary findings in the literature are reconciled by showing that, all else equal, the afternoon is the best time of day for learning, but academic fatigue wears on students as the school day progresses.

Setting:

Description of the research location.

> Data for this study come from the United States Air Force Academy (USAFA). USAFA is a fully accredited post-secondary institution with annual enrollment of approximately 4,500 students, offering 32 majors within the humanities, social sciences, basic sciences, and engineering. The school day at USAFA is highly structured, which is atypical of most universities, but very similar to a high school setting. There are four 53 minute class periods

each morning and three each afternoon after an 85 minute lunch break. All students are required to attend a mandatory breakfast 25 minutes before first period. In this study, we exploit give important features of the school day structure at USAFA. First, students in their freshman year at USAFA are required to take a series of core course in which attendance in their assigned section is mandatory. Second, students are randomly assigned to course sections and cannot choose which periods they take their classes or with which instructors. Third, students are not assigned a class for every class period. Fourth, we exploit the fact that USAFA runs on an M/T schedule. On M days, students have one set of classes and on T days they have a different set of classes. The M/T schedule runs every other day. Thus, the same student has two different class schedules within the same semester. Fifth, perhaps most importantly, all exams for a course are administered to all sections of students during a common testing period and then graded on a common scale. This makes comparisons of student performance in the same course, but different sections unbiased.

The data is an administrative data set from The United States Air Force Academy. It is managed in collaboration with Lt. Col. Scott Carrell and Jim West. The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the USAF, DoD, or the U.S. Government.

Population / Participants / Subjects:
Description of the participants in the study: who, how many, key features, or characteristics.

> Data comes from United States Air Force Academy (USAFA), covering the 2004-2009 school years. I limit the data to freshman in their first semester at USAFA because, despite being a higher-education setting, the schedule of USAFA students is similar in its structure and course load to that of secondary students. Restricting the data only to freshman means I am examining students who developmentally are still very similar to high school students. In total, I examine 4816 students who received 16,119 grades in courses taken. Courses are taught in small sections that average 17 students and required courses may have as many as 12 sections taught in a single semester.

The unit of observation will be an individual student's performance, measured as a normalized grade in a given course. Each observation in my data includes a student id, course, normalized grade, teacher, course section, schedule day, period of the day as well as extensive student background characteristics. I add a number of key constructed variables. Peer characteristics for each section are included. I construct variables concerning the overall load of a student's schedule-day including *Credit Hours* (total number of credits for a schedule-day), *Cumulative Classes* (number of classes up to that point of the day), C*onsecutive Classes* (number of classes in a row) and five "Lead-up" indicator variables that represent what a student had scheduled before the observed class.

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.

> Causal inference is done using linear regressions. The general regression specification is given by:

$$Grade_{icjspt} = \alpha + Period_p + Load_{icspt} + LeadUp_{icspt} + \beta X_i + Peers_{-icspt} + \lambda_{ct} + \psi_{jt} + \rho_i + \epsilon_{icjspt}$$

Where *Grade* is student *i*'s grade in course *c*, taught by professor *j*, on schedule day *s*, class period *p*, in year *t*. *Period* represent a set of fixed effects for each period. *Load* represents the three indicator of a student's course load on a schedule-day. *LeadUp* are indicators of the five lead-up possibilities, *X* is a vector of individual background characteristics and *Peers* represents the peer characteristics (excluding student *i*) in that section. Lastly λ, φ, and ρ represent course, professor and individual student fixed effects, respectively. Because of the separate "M" and "T" schedule-day setup, I am able to look at specifications using individual fixed effects where the identifying variation comes from how a student's performance differs between their two different schedules.

Findings / Results:

Description of the main findings with specific details.

> The above specification provides numerous permutations and Table 2 in the appendix shows early results of specifications that include *Period* and *Load* variables. Period 1 is the omitted period. The coefficients on Periods 2-7 are positive and significant showing that the first, earliest, class of the day is associated with significantly worse performance. 6th and 7th periods, all else being equal, are the best time of day to have a class, they cause an average improvement of around 0.25 of a standard deviation. Of course, later classes are also associated with more cumulative classes. The coefficients on the *cumulative* and *consecutive* variables are negative showing that academic fatigue sets in as the day goes on. These results reconcile findings from physiologists that teenagers are most alert in the late afternoon with economic studies of students showing that productivity declines during the day.

The inclusion of the *Lead Up* variables reveals an interesting pattern with respect to P.E. Show in the bar graph in the appendix, a student in a 3rd period class who just had P.E. will, on average, perform .14 of a standard deviation better than a student who just had two classes. However, take the same situation right before 7th period, the last one of the day, and P.E. is predicted to harm performance by .08 of a standard deviation. It seems that later in the day, students are prone to "checking out" after P.E., while in the morning it provides a boost to academic performance.

A third set of early results come from stratifying the data. When I compare STEM classes to Non-STEM academic classes, it's clear that STEM classes are more impacted by the day's schedule than are Non-STEM. This may be due to differences in teaching structure (lecture vs discussion) or more weight put on out-of-class essays and homework rather than exams. Either way, this suggests that putting STEM classes in the middle of the day, the most productive time, can raise the overall level of learning without. When data are stratified by student's strengths (e.g. students with strong math aptitude vs weaker aptitude), results shed further light on how the day should be structured. Stronger student's grades show more resiliency to time-of-day, academic fatigue and lead-up effects while the coefficients on weaker students tend to amplify. Weak math students are the most negatively impacted by having math first thing in the morning. These results give causal support to schools scheduling a student's weakest subjects before lunch (after which when fatigue kicks in), but after the early morning courses. Students in every school are bound to have different strengths and weaknesses. Intelligently scheduling targeting when during the day the weaknesses are taught could improve student performance without significant impact on the students' strengths.

Conclusions:
Description of conclusions, recommendations, and limitations based on findings.

> Findings from this study shed light on a number elements within a student's daily schedule that can affect their academic achievement. Two similar students taking the same classes with the same teachers, but with different schedules could be expected to get grades as different as two-tenths of a standard deviation (approximately a B- to a B+ difference) in certain scenarios. With each finding, we discuss administrative action that could be taken to better optimize the school schedule. Schools face multitudes of different constraints when it comes to scheduling, but many of our suggestions are quite broad and could be achieved using standard scheduling software.

Our findings regarding physical education further support the idea of shifting STEM classes away from the morning. We find positive effects, which are especially strong for low-ability students, of having P.E. in the morning. This is in line with Lambourne and Tomporowski's (2010) review of studies that have explored the effects of both overall physical fitness and acute exercise. Once again this suggests shifting away (especially for lower-ability students) from morning STEM classes and towards morning physical education to wake students up. Some of our highest point-estimates are of the negative impact of having P.E. just before the last period of the day. We interpret this as a sort of mental “checking-out" whereby it is hard for students to reengage their minds for a final class late in the afternoon. Depending on the facilities at a school, P.E. could be a class where more students could be active in the morning by increasing morning class sizes without requiring more instructors or sections offered.

Academic fatigue, measured by both total and consecutive classes seems to adversely affect performance. Spacing out free periods can help reduce fatigue. Now, it's hardly reasonable for us to advocate later start times (or not having a first period class) in addition to then having multiple spaced-out free periods during the day. High school students are often taking six, sometimes seven classes. However, one takeaway is that free periods during the last period of the day are effectively a waste. They don't help with start time or fatigue. Sports commonly dictate that students have their last period free because of scheduling conflicts, but our evidence suggests that giving students their last period free should be avoided whenever possible.

While most of our discussion has been around what administrators or schedulers could do to better optimize the school schedule, there is also the potential for student optimization. In a setting where a student has some control over which classes they take, they may be able to do themselves a favor by forcing themselves to be active in the morning, spacing out their breaks and not taking too many consecutive classes.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Carrell, S. E., Maghakian, T., & West, J. E. (2011). A's from Zzzz's? The Causal Effect of School Start Time on the Academic Achievement of Adolescents. *American Economic Journal: Economic Policy*, *3*(3), 62–81. doi:10.1257/pol.3.3.62

Cortes, K. E., Bricker, J., & Rohlfs, C. (2012). The Role of Specific Subjects in Education Production Functions: Evidence from Morning Classes in Chicago Public High Schools. *The B.E. Journal of Economic Analysis & Policy*, *12*(1). Retrieved from http://www.degruyter.com/view/j/bejeap.2012.12.issue-1/1935-1682.2749/1935-1682.2749.xml?format=INT

Crowley, S. J., Acebo, C., & Carskadon, M. A. (2007). Sleep, circadian rhythms, and delayed phase in adolescence. *Sleep Medicine*, *8*(6), 602–612. doi:10.1016/j.sleep.2006.12.002

Dills, A. K., & Hernández-Julián, R. (2008). Course scheduling and academic performance. *Economics of Education Review*, *27*(6), 646–654. doi:10.1016/j.econedurev.2007.08.001

Edwards, F. (2012). Early to rise? The effect of daily start times on academic performance. *Economics of Education Review*, *31*(6), 970–983. doi:10.1016/j.econedurev.2012.07.006

Lambourne, K., & Tomporowski, P. (2010). The effect of exercise-induced arousal on cognitive task performance: a meta-regression analysis. *Brain Research*, *1341*, 12–24. doi:10.1016/j.brainres.2010.03.091

Pope, N. (2014). How Time of day affects Productivitiy: Evidence from School Schedules. *Working Paper*.

Appendix B. Tables and Figures

Not included in page count.

Figure 1 – Example USAFA Schedule

Day	Course	Description	Hours	Period Section	Room	Instructor
M	History 100	INTRO TO MILITARY HISTORY	3.00	M1B	5D37	Lt Col Marsha B Ivins
	Math 141	CALCULUS I	3.00	M2D	5D12	Dr. Gregory A Jarvis
	ReadSkls 103	READING ENHANC./4TH CL	2.00	M4A	1A78	Dr. Kathryn Hire
	Russian 131	BASIC RUSSIAN	3.00	M5A	4H18	Ms. Millie Hughes-Fulford
	English 111	INTRO/COMPOSITION & RESEARCH	3.00	M6B	4D6	Capt Charles Hobaugh
	ExtProg 917	INTRAMURALS/GROUPS 1/2	0.00	M7A		
T	Chem 100	APPLICATIONS OF CHEMISTRY I	3.00	T1B T2B	2M117	Capt Kevin B Chilton
	PhyEd 110D	BOXING	0.50	T3A T4A		
	Russian 131	BASIC RUSSIAN	3.00	T5A	4H18	Ms. Millie Hughes-Fulford
U	FYE 101B	FIRST YEAR EXPERIENCE	1.00	U1T		Lt Col Scott L Tingle
	Cmsng Edu 100	4CL COMMISSIONING EDUCATION	0.00	U1A		

Table 1 – Summary Stats of Section Data

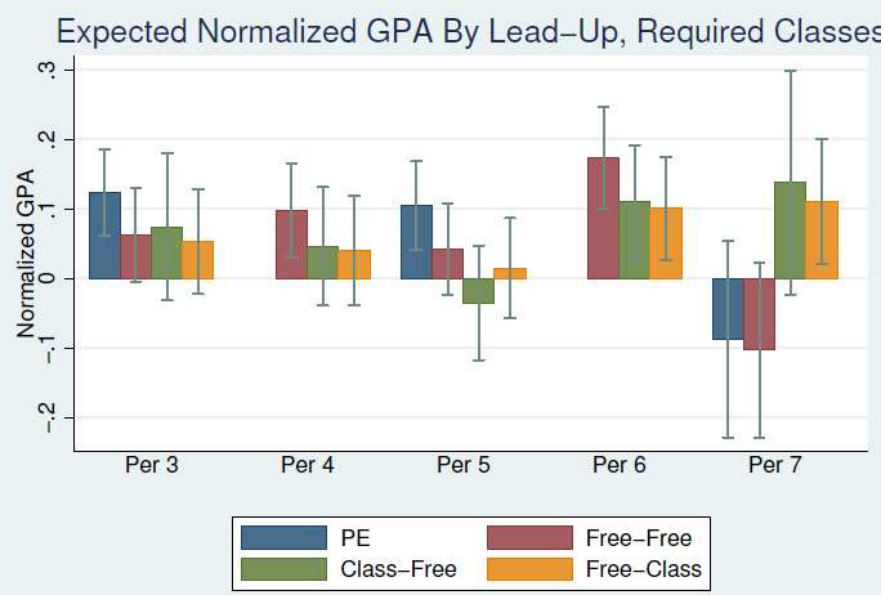
	(1) All Cohorts
Number of Students	4816
Number of Sections	1056
Number of Courses	32
Avg Section Size	16.56
Period 1 Avg	-0.0249
Period 2 Avg	0.117
Period 3 Avg	0.0686
Period 4 Avg	0.0187
Period 5 Avg	0.0718
Period 6 Avg	0.160
Period 7 Avg	0.0568

Table 2 - Results of Period analysis

	(4) Grade	(5) Grade	(6) Grade
2nd Period	0.132***	0.0998***	0.143***
	(0.0280)	(0.0350)	(0.0370)
3rd Period	0.140***	0.127***	0.157***
	(0.0276)	(0.0346)	(0.0363)
4th Period	0.146***	0.140***	0.219***
	(0.0333)	(0.0433)	(0.0466)
5th Period	0.112**	0.132**	0.220***
	(0.0434)	(0.0559)	(0.0617)
6th Period	0.223***	0.276***	0.357***
	(0.0505)	(0.0653)	(0.0739)
7th Period	0.197***	0.244***	0.333***
	(0.0562)	(0.0731)	(0.0857)
Credits/Day	0.00245	0.00269	-0.00392
	(0.00377)	(0.00383)	(0.00385)
Consecutive Classes	-0.0243*	-0.0341**	-0.00931
	(0.0147)	(0.0148)	(0.0157)
Cumulative Classes	-0.0451***	-0.0370***	-0.0387**
	(0.0132)	(0.0136)	(0.0159)
Teacher FEs	No	Yes	Yes
Indv FEs	No	No	Yes
N	16119	16119	16119
R-Square	0.191	0.251	0.717

Standard Errors in parentheses, errors clustered at the section level

Figure 2 - Bar Graph of "Lead-Up" Coefficients



Bars show 90% CI. Having 2 classes (Class-Class) is the omitted category.